YOLOv2 and Hierarchical Soft-Max Tree for US Traffic Sign Detection System

Dat Nguyen
University of California, Davis

PROBLEM STATEMENT

- A system can detect traffic sign with high accuracy and run on real-time (>30 fps).
- Ability to infer to its abstract type when seeing an unknown sign.

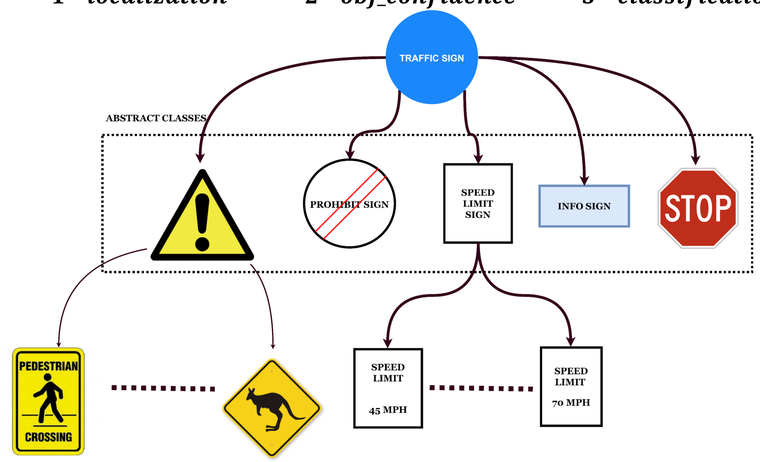
Approach

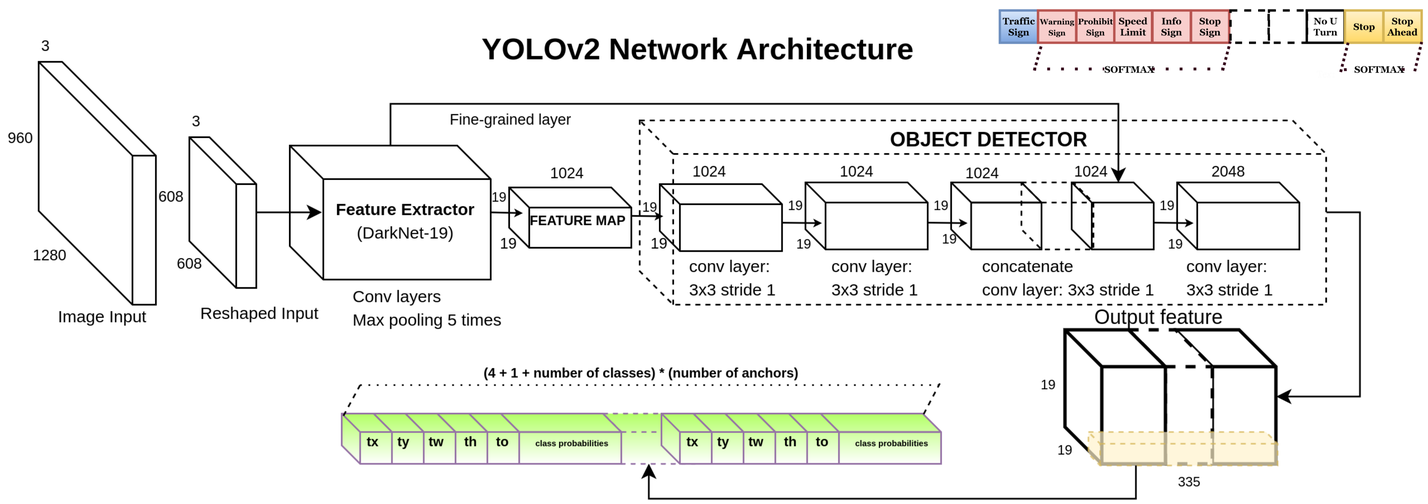
- Built a You Only Look Once version 2 (YOLOv2) for Traffic Sign Detection.
- Embedded prior knowledge of the dataset into the network through a hierarchical graph.

Methods

- Based on the standard YOLO-v2, we extended the network architecture by using light-weight CNN with higher input resolution [2] using Densely Connected [3] and Depth-wise Separable Conv. Layers [4].
- Objective Function :

$$w_1\mathcal{L}_{localization} + w_2\mathcal{L}_{obj\_confidence} + w_3\mathcal{L}_{classification}$$





EXPERIMENTS

- Trained on both LISA + LISA Extension, open-sourced US Traffic Sign Detection Dataset.
- Designed a Soft-Max Tree for joining multiple Traffic Sign dataset

RESULTS

Network	Resolution	Parameters	mAP (*)	Forward Pass(**)
Original YOLOv2	608x608	50.8 M	76.8	36ms
Mobile-YOLOv2	960x960	6.78 M	82.8	**20ms**
Densely-YOLOv2	960x960	10.6 M	**90.8**	55ms

(*) mAP is calculated on actual labels, without abstract classes. Also, this result is preliminary. Additional training might increase the result.
(**) Validated on NVIDIA GTX 1080.

FUTURE WORK

- Use Recurrent Neural Network to take advantage of temporal information.
- Train on larger dataset.
- Occlusion Detection.

ACKNOWLEDGEMENT

Thank you Dr. Christoph Mertz, Research Associate Jina Wang, and all RISS staff for providing extraordinary support, resources and mentorship to help me complete this project.

Carnegie Mellon University
Robotics Institute

References

[1] Redmon, Joseph and Ali Farhadi. "YOLO9000: Better, Faster, Stronger." *CVPR* 2017.
[2] Ashraf, Khalid et al. "Shallow Networks for High-Accuracy Road Object-Detection." *VEHITS* (2017).
[3] Huang, Gao et al. "Densely Connected Convolutional Networks." CVPR 2017.
[4] Howard, Andrew G. et al. "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications." 2017.